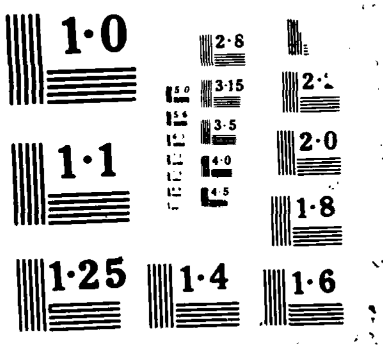
AD-A191 160 ON THE CAPACITY OF CHANNELS WITH UNKNOWN INTERFERENCE 1/1
(U) MICHIGAN UNIV ANN ARBOR COMMUNICATIONS AND SIGNAL
PROCESSING LAB M V HEGDE ET AL. 14 AUG 87
UNCLASSIFIED N00014-85-K-0545 F/G 12/9 NL


1·0
2·8
3·15
3·5
2·0
1·1
4·0
4·5
1·8
1·25
1·4
1·6

(2)

ON THE CAPACITY OF CHANNELS WITH UNKNOWN INTERFERENCE

M. V. Hegde, W. E. Stark, D. Teneketzis

COMMUNICATIONS & SIGNAL PROCESSING LABORATORY
Department of Electrical Engineering and Computer Science
The University of Michigan
Ann Arbor, Michigan 48109

August 1987

Technical Report No. 251
Approved for public release; distribution unlimited.

Prepared for
OFFICE OF NAVAL RESEARCH
Department of the Navy
Arlington, Virginia 22217

and

NATIONAL SCIENCE FOUNDATION
Washington, D.C. 20550

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE

REPORT DOCUMENTATION PAGE

1a. REPORT SECURITY CLASSIFICATION
UNCLASSIFIED

1b. RESTRICTIVE MARKINGS
NONE

2a. SECURITY CLASSIFICATION AUTHORITY

2b. DECLASSIFICATION / DOWNGRADING SCHEDULE

3. DISTRIBUTION / AVAILABILITY OF REPORT
Approved for Public Release;
Distribution Unlimited

4. PERFORMING ORGANIZATION REPORT NUMBER(S)
TR 251

5. MONITORING ORGANIZATION REPORT NUMBER(S)

6a. NAME OF PERFORMING ORGANIZATION
Communications & Signal Processing Laboratory

6b. OFFICE SYMBOL *(If applicable)*

7a. NAME OF MONITORING ORGANIZATION
Office of Naval Research and National Science Foundation

6c. ADDRESS *(City, State, and ZIP Code)*
The University of Michigan
Ann Arbor, Michigan 48109-2122

7b. ADDRESS *(City, State, and ZIP Code)*
ONR, 800 N. Quincy St., Arlington, VA 22217
NSF, Washington, D. C. 20550

8a. NAME OF FUNDING / SPONSORING ORGANIZATION

8b. OFFICE SYMBOL *(If applicable)*

9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER
ONR Contract No. N00014-85-K-0545
NSF Grant No. ECS-8517708

8c. ADDRESS *(City, State, and ZIP Code)*

10. SOURCE OF FUNDING NUMBERS

PROGRAM ELEMENT NO.	PROJECT NO.	TASK NO.	WORK UNIT ACCESSION NO.

11. TITLE *(Include Security Classification)*
On the Capacity of Channels With Unknown Interference

12. PERSONAL AUTHOR(S)
M. V. Hedge, W. E. Stark, D. Teneketzis

13a. TYPE OF REPORT
Tech. Report

13b. TIME COVERED
FROM ________ TO ________

14. DATE OF REPORT *(Year, Month, Day)*
August 14, 1987

15. PAGE COUNT
47

16. SUPPLEMENTARY NOTATION

17. COSATI CODES

FIELD	GROUP	SUB-GROUP

18. SUBJECT TERMS *(Continue on reverse if necessary and identify by block number)*
Communications
Channel Capacity
Jamming Strategy

19. ABSTRACT *(Continue on reverse if necessary and identify by block number)*

We model the process of communicating in the presence of interference, which is unknown or hostile, as a two-person zero sum game with the communicator and the jammer as the players. The objective function we consider is the mutual information. The communicator's strategies are distributions on the input alphabet and on a set of quantizers. The jammer's strategies are distributions on the noise power subject to certain constraints. We consider various conditions on the jammer's strategy set and on the communicator's knowledge. For the case with the decoder uninformed of the actual quantizer chosen we show that, from the communicator's perspective, the worst-case jamming strategy is a distribution concentrated at a finite number of points thereby converting a functional optimization problem into a non-linear programming problem. Moreover, we are able to characterize the worst-case distributions by means of necessary and sufficient conditions which are easy to verify. For the case with the decoder informed of the actual quantizer chosen we are able to demonstrate the existence of saddle-point strategies. The analysis is also seen to be valid for a number of situations where the jammer is adaptive.

20. DISTRIBUTION / AVAILABILITY OF ABSTRACT
☒ UNCLASSIFIED/UNLIMITED ☐ SAME AS RPT. ☐ DTIC USERS

21. ABSTRACT SECURITY CLASSIFICATION
UNCLASSIFIED

22a. NAME OF RESPONSIBLE INDIVIDUAL
Carol S. Van Aken

22b. TELEPHONE *(Include Area Code)*
(313) 764-5220

22c. OFFICE SYMBOL

DD FORM 1473, 84 MAR 83 APR edition may be used until exhausted. All other editions are obsolete.

SECURITY CLASSIFICATION OF THIS PAGE
UNCLASSIFIED

On The Capacity Of Channels With Unknown Interference

M.V. Hegde W.E. Stark*
D. Teneketzis†
Department of Electrical Engineering and Computer Science
University of Michigan
Ann Arbor, MI 48109

August 14, 1987

Abstract

We model the process of communicating in the presence of interference, which is unknown or hostile, as a two-person zero sum game with the communicator and the jammer as the players. The objective function we consider is the mutual information. The communicator's strategies are distributions on the input alphabet and on a set of quantizers. The jammer's strategies are distributions on the noise power subject to certain constraints. We consider various conditions on the jammer's strategy set and on the communicator's knowledge. For the case with the decoder uninformed of the actual quantizer chosen we show that, from the communicator's perspective, the worst-case jamming strategy is a distribution concentrated at a finite number of points thereby converting a functional optimization problem into a non-linear programming problem. Moreover, we are able to characterize the worst-case distributions by means of necessary and sufficient conditions which are easy to verify. For the case with the decoder informed of the actual quantizer chosen we are able to demonstrate the existence of saddle-point strategies. The analysis is also seen to be valid for a number of situations where the jammer is adaptive.

*The work of M.V. Hegde and W.E. Stark was supported by the Office of Naval Research under contract N00014 85 K0545

†The work of D. Teneketzis was supported by the National Science Foundation under Grant no. ECS-8517708 and by a Rackham Research Grant of the University of Michigan

1 Introduction

The applicability of game-theoretic models in jamming situations is by now well established [Blac 57], [Root 61], [McEl 83a], [McEl 83b], [Star 82], [Chan 85], [Peng 86]. In this paper we formulate fairly general models for a number of jamming situations as two-person zero-sum games between the communicator and the jammer. We allow the jammer the choice of one of a set of noise distributions satisfying peak and average power constraints. By way of counter-measure the communicator is allowed to randomize the input symbols as well as randomize the quantizer at the output side. We intend the analysis to be applicable to the performance of soft decision decoding for jammed channels.

Typically in a spread spectrum channel the performance in additive white Gaussian noise is identical to the performance of non-spread systems; namely the bit error probability decreases exponentially with signal-to-noise ratio. However, when subject to worst-case partial-band or pulsed jamming (wherein power is concentrated in time or frequency to affect only a fraction of the symbols transmitted while allowing the remaining to be received "error-free") the bit error probability of a spread-spectrum system decreases only inverse linearly with the signal-to-noise ratio. This is a significant degradation, typically of the order of 30-40 dB for a bit error probability on the order of 10^{-5}.

To remedy this situation most systems use some form of error-correction coding. For example, it can be shown that with a hard decision decoder if the code rate is small ($< 1/2$) and the jammer is allowed to pulse between several Gaussian distributions then there is no loss in signal-to-noise ratio necessary for reliable communications compared to an additive Gaussian noise channel with the same (average) power. So it can be said that coding (with hard decision demodulation) neutralizes a (power constrained) jammer (i.e., makes the performance the same as an additive white Gaussian noise)[Stark 85a],[Ma 84]. It can also be shown that the worst case jamming strategy is to pulse between two zero mean Gaussian noise distributions, one of which has zero variance.

As has been well known in the communication field, hard decision decoding loses roughly 2 dB in signal-to-noise ratio compared to soft decision decoding. Thus considerable interest has focused on soft-decision decoding. One problem that has been observed is that if a (soft) decoding algorithm designed for a non-jammed channel is used for a jammed channel then the performance is extremely poor when the jamming strategy is optimized. One method for "overcoming" this difficulty is to assume the jamming noise is one of two distributions (usually one having zero variance called the "off" state and the other called the "on" state) and that the decoder knows when the jammer is "on" and when the jammer is "off". Using this side information, similar results to the hard decision case have been obtained for the soft decision case [Simo 85] (for small rates there is no loss in performance). However assuming this information is available is assuming away the problem. Most systems analyses do not incorporate jamming strategies that affect the reliability of the side information.

Thus there has been considerable interest over the last few years on decoding algorithms that do not assume side information and do not do hard decision decoding. However, most of these algorithms still assume the jammer pulses between one of two levels. In this paper we investigate the case of a decoder that processes symbols from a finite alphabet and where the only constraints on the jammer are average and peak power. We formulate the problem as a game with two players. The jammer whose strategy set consists of distributions on the power of the jamming noise, and the communicator, whose strategy set consists of a pair of distributions, one on the input alphabet and one on the set of quantizers. We look for worst-case jamming strategies and investigate when the game admits of a saddle point. We do the analysis using mutual information as our objective function.

We consider a modulator that transmits one out of M signals. This transmitted signal is denoted by the random variable X. The received signal which has been corrupted by the jammer in some fashion is demodulated and quantized into one of L values. In order to disallow the jammer from using knowledge of the quantizer in designing his worst-case strategy, we allow randomization of the quantizer over some given set of quantizers. Clearly such randomization increases the the size of the communicator's strategy set. Thus, we view this situation as a game with two players; the jammer and the communicator. The jammer selects the noise in the channel and the communicator chooses the encoder, the decoder and the quantizer. The strategy set for the jammer is the set of all distributions on the power of the jamming noise subject to the given constraints on the peak and average power. The strategy set for the communicator is the set of all distributions on the input alphabet and on the set of quantizers.

For this general set up we show that the worst case jamming strategy from the communicator's perspective is to pulse between a finite number of power levels. We also consider the case of random decoding strategies where the demodulator output is quantized into a finite number of outputs by a randomized quantizer, i.e., the quantization thresholds are random.

For this case we show that the optimal randomized quantizer can perform better than the nonrandomized quantizer and that from the jammer's point of view the worst-case distribution of the thresholds is concentrated on a finite number of points. Our basic model can be easily seen to fit a frequency-hop communication system in which the modulation utilizes an M-ary signal set, which in many cases are orthogonal signals. The spread-spectrum bandwith is divided into a large number of frequency slots. Each possible modulated signal is hopped from frequency slot to frequency slot using a pseudo-random hopping pattern. During each hop one of the M signals is transmitted. There are two important special cases. First, all modulated signals use the same hopping pattern and second, each signal has its own hopping pattern. The demodulator is a coherent or noncoherent matched filter which is then quantized to a finite number of values.

The remainder of the paper is organized as follows. In Section 2 we define the models we will be considering and give examples for which our models apply. In Sections 3 and 4 we derive our results concerning the worst case jamming strategy and the optimal quantizer strategy for the cases with decoder uninformed about the actual quantizer chosen and with decoder informed about the actual quantizer chosen respectively. Finally in Section 5 we discuss our results and state our conclusions and extensions.

2 Channel Models

In this section we describe the models we use in the subsequent analysis. In all cases we consider a modulator that transmits one out of M signals in D dimensions ($D \leq M$). This transmitted signal is denoted by the random variable X. The received signal which is corrupted by the jammer in some fashion is demodulated and quantized into one of L values. The received signal is denoted by the random variable Y. (Y can also be a random vector without changing any of the following analysis).

The general philosophy that we will use is that of game theory with the players being the jammer and the communicator. The jamming strategies are distributions dF on D random variables, $Z_1, Z_2, ..., Z_D$. These random variables represent the power of the jammer in each of the signal dimensions and are modelled as modulating a generic noise variable present in the channel. For example, if $D =$ 1 and N is a zero-mean, variance 1 Gaussian random variable then the jammer's noise may be of the form $Z_1 N$. The jammer, however, has an average power constraint and a peak power constraint. More generally the jammer is constrained by

$$\int f(z_1, z_2, ..., z_D) dF(z_1, z_2, ..., z_D) \ \leq \ K_J \tag{1}$$

and

$$0 \ \leq \ Z_j \ \leq \ b_j \ \ j = 1, \ldots, D \tag{2}$$

where b_j is the peak power constraint and $f(z_1, ..., z_D)$ is some continuous functional of $(z_1, ..., z_D)$. For average power constrained channels with no peak constraint we let b_j become very large.

The output of the demodulator is quantized into one of L values, say $0, 1, ..., L-1$. The output of the quantizer, Y, is also the output of the channel for coding. The strategies for the communicator are to choose a distribution, $dG(\theta)$, on the quantization thresholds and a distribution, $dP(x)$, on the input alphabet. We will let Θ parametrize the quantizers and assume Θ is some compact subset of R (θ will be used to denote both the random variable as well as a particular realization). For each $(z_1, ..., z_D)$ and θ there is a probability distribution on the output of the channel given the input of the channel:

$$Prob\{Y = y | X = x, \Theta = \theta, Z_1 = z_1, Z_2 = z_2, ..., Z_D = z_D\} = p(y|x, \theta, z_1, z_2, ..., z_D). \tag{3}$$

The above model describes the input output relation of the channel for a particular symbol. In addition we model the channel as being memoryless.

We now introduce some notation. Let:

$A \ = \ \{0, 1, ..., M-1\}$ be the input alphabet,

$B = \{0, 1, ..., L-1\}$ be the output alphabet,

Θ be the quantizer parameter space (some compact subset of R)

Z be $(Z_1, ..., Z_D)$, $\quad (0 \leq Z_i \leq b_i)$

$p(y|x, \theta, z)$, the transition probability from x to y given θ, z, and

$P_{yx}(\theta, z)$ the corresponding stochastic matrix, $P_{yx}(\theta, z) = [p(y|x, \theta, z)]$.

We assume that

(i) $p(y|x, \theta, z)$ is continuous in z for all θ, x and

(ii) $p(y|x, \theta, z)$ is continuous in θ for all x, z.

Let **S** denote the set of all probability distributions on the Borel sets of $K \triangleq \{\underline{z} = (z_1, \ldots, z_D) : 0 \leq z_i \leq b_i\}$, and

$$\begin{aligned}
I(G, P; F) &= I\left(\int_K \int_\Theta P_{yx}(\theta, z) dG(\theta) dF(z)\right) \\
&= I\left(\int_K P_{yx}^G(z) dF(z)\right) \\
&= I\left(\int_\Theta P_{yx}^F(\theta) dG(\theta)\right) \\
&= I\left(\overline{P}_{yx}(G, F)\right) \qquad (4)
\end{aligned}$$

where $I\left(\overline{P}_{yx}(G, F)\right)$ is the mutual information whenever X and Y are related by the stochastic matrix $\overline{P}_{yx}$.

The performance measure we are interested in is the largest rate such that nearly error-free communication can be achieved, i.e. channel capacity. Another performance of interest is the channel cutoff rate, R_0, since many researchers believe this to be a practical limit to the set of rates for which reliable communication is possible. Similar results to those in this paper can be derived with R_0 as the performance measure (see [Hegd 87]).

We consider two different information structures for the communicator:

I. The decoder is unaware of the actual quantizer chosen but only knows the distribution $dG(\theta)$ on the set of quantizers. The jammer knows only the set of quantizers but not the distribution $dG(\theta)$ chosen by the communicator. He is also aware that the decoder does not know the actual quantizer chosen.

II. The decoder knows the actual quantizer chosen. Again the jammer knows only the the set of quantizers. He also knows that the decoder is aware of the actual quantizer chosen.

Case I is seen to apply to situations where, for reasons of implementation perhaps, the decoding is fixed and not altered with the specific quantizer chosen. It may also be viewed as worst-case in the sense that the decoder's knowledge of the specific quantizer and the utilization of such knowledge can only improve the communicator's performance. When there is no randomization of the quantizer, i.e. the quantizer is fixed, Cases I and II are the same and our results for both cases apply to that situation. Also several special jamming strategies are of interest because of correspondence with physical problems. We will classify the cases as follows.

A. Arbitrary joint distribution on $Z_1, Z_2, ..., Z_D$.

B. $Z_1 = Z_2 = ... = Z_D = Z$.

C. One dimensional jamming, i.e., at most one of the random variables $Z_i \neq 0$.

D. Independent jamming, i.e., $Z_1, Z_2, ..., Z_D$ are independent.

Case B corresponds to the physical situation where the jammer is not able to place different amounts of power in different dimensions of the signal space. Case C corresponds to the case where only one of the dimensions can be jammed at once. Case D corresponds to a frequency-hop communication system with independent hopping for the different symbols. The standard game theoretic description is given below.

Communicator's Perspective

The communicator is interested in the maximum rate at which information can be reliably transmitted no matter what strategy the jammer employs. The communicator designs his system assuming the jammer will somehow find out the strategy he is using and then choose the worst possible distribution on the power levels. In Case I the largest rate for which information can reliably be transmitted is

$$\max_{G,P} \min_{F} I(G, P; F)$$

where $I(G, P; F) \triangleq I(X; Y)$ when (dG, dP) is chosen by the communicator and dF is chosen by the jammer. That this is the maximum rate of reliable transmission is well known since what we are dealing with is a compound channel with a finite input alphabet and a finite output alphabet [Csiz 81, pgs. 172-173].

Jammer's Perspective

The jammer is interested in the minimum rate such that information can not be reliably transmitted at any higher rate no matter what strategy the communicator employs. The jammer designs his system assuming the communicator will somehow find out the strategy he is using and then design the optimal communication system. In Case I the smallest rate that the jammer can guarantee reliable communication can not be above is

$$\min_{dF} \max_{dG,dP} I(G,P;F).$$

That this is the smallest rate the jammer can guarantee is obvious since for each F the rate above which reliable communication is impossible is $\max_{dG,dP} I(G,P;F)$. In case II the appropriate mutual information can be written as an expectation of the mutual information for a fixed θ:

$$I(G,P;F) = E_G(I(\theta,P;F))$$

where E_G refers to taking expectations w.r.t. dG and $I(\theta,P;F) \triangleq I(X;Y|\theta)$.

In all of our analysis we assume that the jammer and the decoder/quantizer have complete information about the set of strategies possible for each other so that no secret information is considered. As mentioned previously, the performance measure we consider is the largest rate such that reliable communication (in the sense of arbitrarily small error probability) is possible. The type of channels we are considering are known as compound channels. We consider the strategies (distributions) by the jammer to be constant for a whole codeword as opposed to (possibly) changing after each symbol of a codeword which would correspond to an arbitrarily varying channel. For compound channels the capacity with finite input and output is well known to be the maximum of the minimum mutual information. The minimum is over all possible transition probabilities and the maximum is over all probability distributions on the input to the channel. Thus, using the maximum of the minimum mutual information as the performance measure corresponds to the largest rate such that reliable communication is possible no matter what strategy the jammer employs. We are now ready to state the results. In brief our results show that when the decoder is informed of the quantization rule then (under a compatibility assumption), there is a saddlepoint in cases A and B, i.e. the jammer's rate and the communicator's rate are equal (Theorem 5). However, when the decoder is not informed of the quantization rule then the jammer's rate and the communicator's rate may differ. However the optimal distributions, F from the communicator's point of view and the G from the jammer's point of view are finite dimensional (in all the cases A, B, C and D) (Theorem 1). This converts a functional optimization problem into a finite-dimensional non-linear programming problem.

3 Case AI: Decoder Uninformed

The communicator has to determine the distributions $(dG(\theta), dP(x))$ that maximize the amount of information $I(G, P; F)$ transmitted. The jammer has to find the noise distribution $dF(z)$ to minimize the information received by the decoder. Thus, the quantizer's goal is to achieve

$$\max_{dG(\theta),dP(x)} \min_{dF(z)} I(G,P;F)$$

whereas the jammer wants to achieve

$$\min_{dF(z)} \max_{dG(\theta),dP(x)} I(G,P;F).$$

In this section we show that for any choice of strategy of either player there is a simple characterization of the optimal reaction strategy of his opponent.

Theorem 1: a) The jammer can achieve the minimum in $\max\limits_{dG(\theta),\ dP(x)} \min\limits_{dF(z)} I(G,P;F)$ with a distribution concentrated at at most $M(L-1)+2$ points.

b) The communicator can achieve the maximum in $\min\limits_{dF(z)} \max\limits_{dG(\theta),\ dP(x)} I(G,P;F)$ with a distribution concentrated at at most $M(L-1)+1$ points.

Discussion: Theorem 1(a) says that the communicator in trying to achieve $\max\limits_{dG(\theta),\ dP(x)} \min\limits_{dF(z)} I(G,P;F)$ has to consider only reaction strategies of the jammer that have a finite number of points of support, i.e. for each $(dG(\theta), dP(x))$ chosen by the communicator the worst-case jammer distribution may be assumed to be concentrated at a finite number of points and this number is bounded uniformly (in $(dG(\theta), dP(x))$) by $M(L-1)+2$. It follows that for a fixed quantizer (i.e. no randomization of the quantization) the worst-case jammer is one who chooses such a finite-dimensional distribution. Similarly Theorem 1(b) says that the jammer may, from his perspective of trying to achieve $\min\limits_{dF(z)} \max\limits_{dG(\theta),\ dP(x)} I(G,P;F)$, consider only finite dimensional reaction strategies on the communicator's part.

To prove these results we use the following facts: (1) the convexity and concavity properties of the mutual information function (it is convex in the channel transition matrix and concave in the input distribution), (2) the equivalence of weak convergence with Levy convergence in our situation [Hegd 87] which we use to show the continuity of our objective function in the strategies as well as compactness of our strategy sets (this allows us to conclude that there is a worst case jamming strategy and a best case communicator strategy) and (3) Dubins' Theorem in order to demonstrate that the optimal reaction strategies are described by distributions concentrated on a finite number of points. Dubins' Theorem allows the extreme points of certain convex sets to be written as finite linear combinations of extreme points of larger convex sets.

Proof of Theorem 1:

We prove part (a) in detail. The modifications required to obtain part (b) are obvious. We start by first proving two intermediate results. Lemmas 1 and 2.

Lemma 1: $I(G, P; F)$ is a Levy-continuous functional of $dF(z)$ for any fixed $(dG(\theta), dP(x))$.

Proof of Lemma 1:

First we note that for every $(dG(\theta), dP(x))$, $I(P_{yx})$ is a convex function of P_{yx} [Csiz 81, pg. 50], i.e.,

$$I(\alpha P_{yx}^1 + (1-\alpha)P_{yx}^2) \leq \alpha I(P_{yx}^1) + (1-\alpha)I(P_{yx}^2) \quad 0 \leq \alpha \leq 1$$

and

$$p(y|x, z) = \int_\Theta p(y|x, \theta, z)\, dG(\theta)$$

is a continuous function of z (since $p(y|x,\theta,z)$ is continuous in z and $p(y|x,\theta,z) \leq 1$, this follows from the Dominated Convergence Theorem). Also

$$\begin{aligned} p(y|x) &= \int_K \int_\Theta p(y|x,\theta,z)\, dG(\theta)\, dF(z) \\ &= \int_K p(y|x,z)\, dF(z). \end{aligned}$$

Hence $p(y|x)$ is a Levy-continuous functional of $dF(z)$ and therefore P_{yx} is a Levy-continuous functional of $dF(z)$.

Now $I(G,P;F)$ is a convex function of P_{yx} and hence it is continuous in the interior of the finite-dimensional set $\mathbf{W}$ of all stochastic matrices. (Thus, $I(G,P;F)$ is continuous at any point P_{yx} such that at least one row of P_{yx} is not a one point distribution, i.e. P_{yx} is not deterministic). Hence, $I(G,P;F)$ is a Levy-continuous function of $dF(z)$ for any fixed $(dG(\theta), dP(x))$. □

Let $\mathbf{S} \triangleq$ set of all probability distributions on the Borel subsets of K, and

$$\mathbf{S}^1 \triangleq \{dF(z) \in \mathbf{S} : \int f(z)\, dF(z) = K_J\} \tag{5}$$

be a hyperplane in $\mathbf{S}$.

Lemma 2: $I(G,P;F)$ achieves its maximum (minimum) in $\mathbf{S}^1$.

Proof of Lemma 2:

We note that $\mathbf{S}$ is compact in the Levy topology [Hegd 87, Appendix C].

Also $\mathbf{S}^1$ is a hyperplane in $\mathbf{S}$ which is closed (since $dF(z) \to \int_K f(z)\, dF(z)$ is Levy-continuous) in the Levy topology.

Hence $\mathbf{S}^1$ being a closed subset of a compact set is itself (Levy)compact.

Thus Lemma 1 asserts that for fixed $(dG(\theta),\ dP(x)), I(G,P;F)$ is a Levy-continuous functional on the compact set $\mathbf{S}^1$. Hence it achieves its minimum (maximum) at some point $dF^*(z) \in \mathbf{S}^1$. □

The above lemmas are now used to complete the proof of Theorem 1.

From Lemma 2 we know that $I(G,P;F)$ achieves its minimum in $\mathbf{S}^1$. Denote the corresponding P_{yx} as $P^*_{yx} = [p^*(y|x)]$ i.e.

$$P^*_{yx} = \int_K \int_\Theta p(y|x,\theta,z)\ dG(\theta)\ dF^*(z). \tag{6}$$

Now consider the set

$$\begin{aligned}\Lambda &= \{dF(z) \in \mathbf{S}^1 : \int_K \int_\Theta p(y|x,z,\theta)\ dG(\theta)\ dF(z) \\ &= p^*(y|x),\quad x \in A, y \in B^1\}\end{aligned} \tag{7}$$

where $B^1 = \{0,1,\ldots L-2\}$. The set Λ is the intersection of $\mathbf{S}$ with $M(L-1)+1$ hyperplanes viz. $\mathbf{S}^1$ and the $M(L-1)$ hyperplanes

$$h_{yx} = \{dF(z)\ \in\ \mathbf{S}^1 : \int_K \int_\Theta p(y|x,z,\theta)\ dG(\theta)\ dF(z)\ =\ p^*(y|x)\}. \tag{8}$$

Furthermore:

$\mathbf{S}$ is convex.

$\mathbf{S}$ is linearly bounded ($\mathbf{S}$ being compact in a metric space is bounded and hence its intersection with any line is bounded).

$\mathbf{S}$ being a compact subset of a metric space is closed and any line l in the metric space is closed. Thus $\mathbf{S}$ is also linearly closed.

Hence we have that $\mathbf{S}$ is a convex, linearly closed and linearly bounded set. By Dubins' Theorem [Dubi 62] we can conclude that since Λ is the intersection of $\mathbf{S}$ with $M(L-1)+1$ hyperplanes, every extreme point of Λ is a convex combination of $M(L-1)+2$ or fewer points of $\mathbf{S}$.

From our construction of Λ we know that $I(G,P;F)$ is constant on Λ. Hence for fixed $(dG(\theta),\ dP(x))$, $I(G,P;F)$ assumes its minimum value at an extreme point of Λ also.

Hence, $I(G,P;F)$ assumes its minimum value at some point $dF(z)$ which is a convex combination of $M(L-1)+2$ or fewer extreme points of $\mathbf{S}$.

Since the extreme points of $\mathbf{S}$ are the one-point distributions, we can finally assert that for each $(dG(\theta), dP(x))$ the jammer can achieve the minimum in

$$\max_{dG(\theta),\ dP(x)} \quad \min_{dF(z)} \quad I(G,P;F)$$

with a distribution concentrated at $M(L-1)+2$ points. This concludes the proof of (a).

For channels which are symmetric for each θ and z i.e. $p(y|x_1, z, \theta)$ is some permutation of $p(y|x_i, z, \theta)$ we see that the set Λ is actually the intersection of $\mathbf{S}$ with $(L-1)+1$ hyperplanes only and hence part(a) of the theorem holds with $(L-1)+2 = L+1$ instead of $M(L-1)+2$. For M-ary symmetric channels, i.e. channels with M inputs and M outputs and such that for each θ and z, $p(y_i|x_i, z, \theta) = 1-\epsilon$ and $p(y_i|x_j, z, \theta) = \dfrac{\epsilon}{M-1}$, $i \neq j$, the bound on the number of points of support reduces to 3.

For (b) we note that the jammer wants to achieve

$$\min_{dF(z)} \max_{dG(\theta), dP(x)} I(G, P; F).$$

This may be written as

$$\min_{dF(z)} \max_{dG(\theta)} C(G, F)$$

where $C(G, F) \triangleq \max_{dP(x)} I(G, P; F)$.

We note that similarly to Lemma 1 for any fixed $dF(z)$, $C(G, F)$ is a continuous functional of $dG(\theta)$. (Simply note that $C(G, F)$ being the maximum of functions convex in P_{yx} is also convex in P_{yx} and proceed as before). Using our hypothesis that $p(y|x, \theta, z)$ is continuous in θ we can show that

$$\min_{dF(z)} \max_{dG(\theta)} C(G, F)$$

can for any $dF(z)$ be achieved by the decoder/quantizer by a distribution $dG(\theta)$ that is concentrated at at most $M(L-1)+1$ points.

Again for symmetric channels we note that part(b) of the theorem holds with L instead of $M(L-1)+1$. For M-ary symmetric channels this number is 2. The number of points of support is one less than Case A as we have not imposed any constraints on the distributions $dG(\theta)$ chosen by the quantizer. □

3.1 Necessary and Sufficient Conditions

We now characterize the aforementioned finite-dimensional distributions by means of necessary and sufficient conditions. We first briefly introduce the necessary definitions and results from optimization theory and then specialize them to our cases.

Let Ω be a convex set and let f be a function from Ω into $\mathbf{R}$. For some fixed x_0 if for all x

$$\lim_{\alpha \downarrow 0} \frac{f((1-\alpha)x_0 + \alpha x) \; - \; f(x_0))}{\alpha} \tag{9}$$

exists f is said to be weakly differentiable at x_0 and the above limit is denoted as $f'_{x_0}(x)$, the weak derivative at x_0. If f is weakly differentiable in Ω at x_0 for all x_0 in Ω, f is said to be weakly differentiable in Ω. We now state an Optimiization Theorem that follows from [Luen 69, pg. 178].

Optimization Theorem: Let f be a continuous, weakly differentiable, convex-cap (concave) map from a compact, convex set to $\mathbf{R}$. Let

$$C \triangleq \sup_{x \in \Omega} f(x). \tag{10}$$

Then

1. $C = \max f(x) = f(x_0)$ for some $x_0 \in \Omega$.

2. A necessary and sufficient condition for $f(x_0) = C$ is $f'_{x_0}(x) \leq 0$ for all $x \in \Omega$.

Constrained Optimization Theorem: [Luen 69, pg. 217] Let Ω be a convex subset of a linear vector space and f and g convex-cap functionals on Ω to $\mathbf{R}$. Assume there is an $x_1 \in \Omega$ such that $g(x_1) < 0$ and let

$$C' \triangleq \sup_{\substack{x \in \Omega \\ g(x) \leq 0}} f(x). \tag{11}$$

If C' is finite then there exists a constant $\lambda \geq 0$ such that

$$C' = \sup_{x \in \Omega} [f(x) - \lambda g(x)]. \tag{12}$$

Furthermore if the supremum in the first equation is achieved by $x_0 \in \Omega$ and $g(x_0) \leq 0$, it is achieved by x_0 in the second equation and $\lambda g(x_0) = 0$. [Luen 69, pg. 217].

Now given any $dG(\theta)$ and the power constraint we define

$$U_c(K_J, G) \triangleq \sup_{\substack{F \in \mathbf{S} \\ h_F \leq K_J}} - I(G, P; F) \tag{13}$$

where $h_F \triangleq \int_K f(z)\, dF(z)$. To simplify notation we define

$$D : \mathbf{S} \rightarrow \mathbf{R} \text{ by } D(F) = \int_K f(z) dF(z) - K_J. \tag{14}$$

Using the Constrained Optimization Theorem we will infer in Theorem 2 that there exists a non-negative constant

$$\lambda = \lambda(G, K_J) \text{ for } D(F) \leq 0 \text{ such that}$$

$$U_c(G, K_J) = \sup_{F \in \mathbf{S}} [-I(G, P; F) - \lambda D(F)]. \tag{15}$$

We now formulate necessary and sufficient conditions for the characterization of the optimal distributions of Theorem 1 in the following two theorems.

Theorem 2: $U_c(G, K_J)$ is achieved by a distribution $F_0 \in \mathbf{S}$ satisfying $D(F) \leq 0$ and a necessary and sufficient condition for $U_c(G, K_J) = -I(G, P; F_0)$ is that for some constant $\lambda \geq 0$

$$\int_K [-i(z; G, F_0) - \lambda f(z)] dF(z) \leq -I(G, P; F_0) - \lambda K_J \tag{16}$$

for all $F \in \mathbf{S}$

where $i(z; G, F_0) \triangleq \sum_{x,y} p(x)\ p(y|x,z) \log \left(\frac{\int p(y|x,z)\ dF_0\ (z)}{\sum_x p(x)\ \int p(y|x,z)\ dF_0(z)} \right)$.

Proof of Theorem 2:

$D : \mathbf{S} \rightarrow \mathbf{R}$ is clearly linear, bounded, convex-cap, continuous and weakly differentiable in $\mathbf{S}$ with $D'_{F_1}\ (F_2) = D(F_2) - D(F_1)$. By choosing F_1 as a distribution with unit mass appropriately we can infer that $D(F_1) < 0$.

Next we show that $I(G, P; F)$ is convex in F.

$$
\begin{aligned}
I(G, P; \alpha F_1 + (1-\alpha)F_2) &= I(\overline{P}_{yx}\ (G, \alpha F_1 + (1-\alpha)F_2)) \\
&= I\left(\int_K \int_\Theta p(y|x,\theta,z)\ dG(\theta)\ (\alpha dF_1 + (1-\alpha) dF_2) \right) \\
&= I(\alpha\ \overline{P}_{yx}\ (G; F_1)\ +\ (1-\alpha)\ \overline{P}_{yx}\ (G; F_2)) \\
&= I(\alpha\ \overline{P}^1_{yx}\ + (1-\alpha)\ \overline{P}^2_{yx}) \\
&\leq \alpha I(\overline{P}^1_{yx})\ +\ (1-\alpha)\ I(\overline{P}^2_{yx}) \\
&\qquad \text{(by the convexity of } I(.)\ w\ r.\ t\ P_{yx}) \\
&= \alpha I(G, P; F_1)\ + (1-\alpha)\ I(G, P; F_2). \qquad (17)
\end{aligned}
$$

Then, since $U_c(G, K_J)$ is finite we can infer from the Constrained Optimization Theorem that there exists some constant $\lambda \geq 0$ such that $U_c = \sup_{F \in \mathbf{S}} [-I(G, P; F) - \lambda D(F)]$.

Now \ 3how that $I(G, P; F)$ is weakly differentiable at all $F \in \mathbf{S}$.

Let $L(\alpha) = I(G, P; \alpha F_1 + (1-\alpha)F_2)$. Since $I(G, P; F)$ is convex in F, $L(\alpha)$ is convex in α. Therefore $\frac{L(\alpha) - L(0)}{\alpha}$ is non-decreasing in α and bounded from below and thus $\lim_{\alpha \downarrow 0} \frac{L(\alpha) - L(0)}{\alpha}$ exists. Furthermore

Lemma 3: $I'_{F_1}\ (G, P; F_2)\ =\ \int\ i(z; G, F_1)\ dF_2(z)\ -\ I(G, P; F_1).$

Proof of Lemma 3:

See Appendix A.

We now have that $I(G, P; F) - \lambda\ D(F)$ is convex-cap, continuous and weakly differentiable in F. Thus, by the Optimization Theorem there is a distribution function $F_0 \in \mathbf{S}$ such that $U_c(G, K_J) = I(G, P; F_0) - \lambda D(F_o)$. The necessary and sufficient condition becomes

$$- I'G, P; F_0{}'\ (F)\ -\ \lambda\ D'_{F_0}\ (F)\ \leq\ 0 \text{ for all } F\ \in\ \mathbf{S} \tag{18}$$

or

$$\int_K [-i(z; G, F_0)\ -\ \lambda f(z)] dF(z)\ \leq\ -I(G, P; F_0) - \lambda h_{F_0}. \tag{19}$$

If $h_{F_0} < K_J$ the power constraint is trivial and the constant λ is zero i.e. $D(F_0) < 0$ but $\lambda\ D\ (F_0) = 0$. Thus the necessary and sufficient condition is established. □

From Theorem 1 we know that it is possible to find F_0 from the set of distributions with a finite number of points of support. Finding such an F_0 entails determining the set of points of increase as well as the amounts of increase of F_0 at those points. Let E_0 denote the set of points of increase of F_0. We now show

Theorem 3: Let F_0 be a probability distribution satisfying the power constraint. Then F_0 achieves $U_c(G, K_J)$ iff for some $\lambda \geq 0$

C1) $-i(z; G, F_0) \leq -I(G, P; F_0) + \lambda(f(z) - K_J)$

for all $z\ \in\ K$.

C2) $-i(z;G,F_0) = -I(G,P;F_0) + \lambda(f(z) - K_J)$

for all $z \in E_0$.

Proof of Theorem 3:

The sufficiency is clear because if both conditions C1 and C2 the conditions of Theorem 2 hold. We show the necessity.

Assume that F_0 is "optimal" but C1 is not true. Then there must exist some $z_1 \in K$ such that $-i(z;G,F_0) > -I(G,P;F_0) + \lambda(f(z) - K_J)$. Let $F_1(z)$ be a probability distribution with a unit increase at such a point $z_1 \in K$. Then

$$\int_K [-i(z;G,F_0) - \lambda f(z)]\, dF_1(z) > -I(G,P;F_0) - \lambda K_J \tag{20}$$

which contradicts Theorem 2. Hence C1 must be true.

Now assume that F_0 is "optimal" but C2 is not true. Then since C1 is true $-i(z;G,F_0) < -I(G,P;F_0) + \lambda(f(z) - K_J)$ for all x in E' where E' is some subset of E_0 with positive measure, i.e.

$$\int_{E'} d\, F_0(z) = c > 0. \tag{21}$$

Since $\int_{E_0 - E'} dF_0(z) = 1 - c$ and on $E_0 - E'$

$$i(z;G,F_0) = I(G,P;F_0) + \lambda(f(z) - K_J) \tag{22}$$

and

$$\int_K [i(z;G,F_0) - \lambda F(z)]\, dF_0(z) = \int_{E'} [i(z;G,F_0) - \lambda f(z)]\, dF_0(z)$$

$$+ \int_{E_0 - E'} [i(z;G,F_0) - \lambda(f(z)]\, dF_0(z)$$

$$\int_{K-E_0} [i(z;G,F_0) - \lambda(f(z)]\, dF_0(z)$$

we have

$$-\, I(G,P;F_0) - \lambda K_J \;<\; -I(G,P;F_0) - \lambda K_J \quad \text{i.e. a contradiction.} \tag{23}$$

Hence C2 must be true too. □

Theorems 1 and 3 reduce the calculation of the distributions describing the reaction strategies to finite-dimensional non-linear programming problems. They can be used to simplify the search for conservative strategies which are optimal for either player. In Theorem 4 below we assert the existence of conservative strategies for each player.

Theorem 4: For the game described in Case AI, there exists a conservative strategy $(d\overline{G}(\theta),\ d\overline{P}(x))$ for the communicator and a conservative strategy $d\overline{F}(z)$ for the jammer, i.e. strategies such that

$$\text{i)} \min_{dF(z)} I(\overline{G},\overline{P};F) \;=\; \max_{dP(x)dG(\theta)} \;\; \min_{dF(z)} \;\; I(G,P;F) \quad \text{and} \tag{24}$$

$$\text{ii)} \max_{dP(x),dG(\theta)} I(G,P;\overline{F}) \;=\; \min_{dF(z)} \;\; \max_{dP(x),dG(\theta)} \;\; I(G,P;F) \tag{25}$$

Proof of Theorem 4:

From Lemmas 1 and 2 we note that

a) $I(G,P;F)$ is lower-semicontinuous in $dF(z)$ for each $(dG(\theta),dP(x))$ and

b) There exists $(dG(\theta), dP(x)) \ni I(G, P; F)$ is lower semi-compact in $dF(z)$.

Theorem 4(i) now follows from a fundamental existence theorem [Aubi 82, pg 209, Th. 1]. Theorem 4(ii) follows similarly. □

3.2 The Remaining Cases

Case BI: With $F(z)$ now recognized as a one-dimensional distribution Theorems 1 and 2 are easily seen to be true.

Case CI: We redefine **S** as follows: $\mathbf{S} = \bigcup_{i=1}^{M} L_i$ where L_i is the space of product distributions such that

$$Pr(Z_i \geq 0) \geq 0$$

$$Pr(Z_j = 0) = 1 \; j \neq i.$$

By previous arguments each L_i is Levy compact and hence so is **S**. Now the proofs of Th. 1 and Th. 2 follow as before.

Case DI: We perform the analysis by fixing $D-1$ of the D distributions $dF_1, \ldots, dF_D$. By minor modifications in the proof of Lemma 1 we see that $I(X;Y)$ is a Levy continuous functional of $dF_i(z)$ for each i. Defining **S** and $\mathbf{S}^1$ similarly except that now both are spaces of distributions of $dF_i(z_i)$ instead of $dF(z)$ we see that for each $(dG(\theta), dP(x))$ the jammer can achieve the minimum in

$$\max_{(dG(\theta),dP(x))} \; \min_{dF(z)=dF_1(z_1),dF_2(z_2),\ldots dF_D(z_D)} I(G,P;F) \tag{26}$$

with a distribution dF_i concentrated at at most $M(L-1)+2$ points.

Since i is arbitrary we can assert that the jammer can achieve the minimum in (16) with distributions dF_i, $i = 1, \ldots, D$ each of which are concentrated at at most $M(L-1)+2$ points. Part (b) of Theorem 1 and Theorem 2 are easily seen to be true as stated for this case.

4 Case AII: Decoder Informed

We have an arbitrary joint distribution on $Z_1, \ldots, Z_D$ the jammer chooses $dF(z)$ and knows that the decoder knows Θ. The communicator chooses $dG(\theta)$ and further the decoder knows θ.

In this case we make a "compatibility" assumption, that is, for every θ and $dF(z)$ the capacity-achieving input distribution $dP(x)$ remains the same.

While "compatibility" certainly restricts our model applicability, we show by example that it is often a worst-case assumption. For instance, we know [Dobr 59] that if $M = L$ and if the jammer's strategy set is restricted such that for each distribution $dF(z)$ and quantizer θ, Prob $\{ \text{error} | x \} \leq \epsilon$ for every x, then the saddle-point strategy for the jammer is to choose a distribution such that

$$p(y|x) = \frac{1}{M} \quad \text{for all } y, x \text{ if } \epsilon > 1 - \tfrac{1}{M}$$

and

$$\begin{aligned} p(y|x) &= \frac{\epsilon}{M-1} \quad y \neq x \text{ if } \epsilon \leq 1 - \tfrac{1}{M} \\ &= 1 - \epsilon \quad y = x \end{aligned}$$

and the saddle-point strategy for the communicator is to choose a uniform distribution on the input alphabet. In our model this corresponds to choosing the canonical noise variables so that $p(y|x,\theta)$ is a symmetric channel for each θ. Such symmetry (and thereby "compatibility") is obtained in a number of other situations as a saddle-point strategy. Under certain conditions, when we have convex constraints in the M noise variables affecting the M inputs of the channel which are invariant under any permutation of the M variables (i.e. a "symmetric" constraint) then the choice of a uniform distribution on the input and the choice of a symmetric channel are saddle-point strategies for the communicator and the jammer respectively (see Appendix B). To describe one more example, if we have M inputs and M outputs,

$$\begin{aligned} y_i &= n_i \quad i = 1, \ldots, M \quad i \neq j \\ y_j &= A + n_j \quad i = j, \end{aligned}$$

n_i are $N(0, v_i), i = 1, \ldots, M$ independent random variables and there is further the constraint $\sum_{i=1}^{M} v_i = c$, then, from arguments similar to those in Appendix B, it can be seen that the saddle point strategy is to choose $v_i = \frac{c}{M}$ and a uniform distribution on the input.

Utilization of the "compatibility" assumption allows us to write the above as

$$\min_{dF(x)} \max_{dG(\theta)} E_G(C(\theta, F)).$$

and

$$\max_{dG(\theta)} \min_{dF(x)} E_G(C(\theta, F))$$

where $C(\theta, F) = \max_{dP(x)} I(\theta; F)$ and $I(\theta; F) = I(X; Y|\theta)$.

In this section we prove the existence of a saddlepoint. The main result is stated in the following theorem:

Theorem 5: There exists a pair of distributions $dG^*(\theta), dF^*(z))$ such that

$$E_G(C(\theta, F^*)) \leq E_{G^*}(C(\theta, F^*)) \leq E_{G^*}(C(\theta, F))$$

for all feasible $dG(\theta), dF(z)$, i.e., $(dG^*(\theta), dF^*(z))$ is a saddle point for the game in case AII.

Proof of Theorem 5: The set of all feasible dF 's i.e.

$$\{dF(z) : \int_K f(z)dF(z) \leq K_J\} \quad 0 \leq z_i \leq b_i$$

is clearly convex and compact. The set of all dG 's is also convex and compact.

We note that for any fixed $dF(z), C(\theta, F)$ is a continuous function of θ.

$$p(y \mid x, \theta) = \int_K p(y|x, \theta, z)dF(z)$$

is by our earlier arguments a continuous function of θ.

Hence, $P_{yx}(\theta)$ is a continous function of θ. Also $C(\theta, F) = C(P_{yx}(\theta))$ and we know that $C(P_{yx}(\theta))$ is convex in $P_{yx}(\theta)$.

Therefore, for every $\theta \in \Theta \ni P_{yx}(\theta)$ is not deterministic, $C(P_{yx}(\theta))$ is a continuous function of $P_{yx}(\theta)$. Hence, for fixed $dF(z), C(\theta, F) = (C(P_{yx}(\theta)))$ is a continuous function of θ and so

$$E_G(C(\theta, F)) = \int_\Theta C(\theta, F)dG(\theta) \tag{27}$$

is a Levy continuous functional of $dG(\theta)$.

Since $E_G(C(\theta, F))$ is linear it is also a concave function of $dG(\theta)$ in dG(θ)). Next we note that $C(\theta, F)$ is convex in $dF(z)$ for each θ since $C(\theta, F) = C(P_{yx}(\theta))$. Hence

$$C(\theta, \alpha F^1 + (1-\alpha)F^2) < \alpha C(\theta, F^1) + (1-\alpha)C(\theta, F^2) \quad 0 \leq \alpha \leq 1.$$

Taking expectations w.r.t. G

$$\int_\Theta C(\theta, \alpha F^1 + (1-\alpha)F^2) dG(\theta)$$

$$\leq \int_\Theta (\alpha C(\theta, F^1) + (1-\alpha)C(\theta, F^2)) dG(\theta)$$

$$\therefore E_G(C(\theta, \alpha F^1 + (1-\alpha)F^2)$$

$$\leq \alpha E_G(C(\theta, F^1)) + (1-\alpha)E_G(C(\theta, F^2)).$$

Consequently, $E_G(C(\theta, F))$ is a convex function in $dF(z)$.

Also $E_G(C(\theta, F))$ is Levy-continuous in $dF(z)$. To prove this it suffices to show that for any sequence F_n converging to F in the Levy metric

$$E_G(C(\theta, F_n)) \to E_G(C(\theta, F)).$$

Since convergence in the Levy metric is in our case equivalent to weak convergence [Hegd 87, Appendix C] it suffices to show this for $F_n \xrightarrow{w} F$. However,

$$\lim_n \; E_G(C(\theta, F_n))$$

$$= \lim_n \int_\Theta C(\theta, F_n) dG$$

$$= \int_\Theta \lim_n C(\theta, F_n) dG \quad \text{(by the Dominated Convergence Theorem)}$$

$$= \int_\Theta C(\theta, F) dG \quad (\text{since } C(\theta, F) \text{ is Levy} - \text{continuous in } F)$$

$$= E_G(C(\theta, F))$$

which proves Levy-continuity in $dF(z)$. From these properties of the objective function and the convexity and compactness of the feasible strategy sets we recognize that the hypotheses of the Sion minmax theorem of game theory are satisfied [Aubi 82, Th7, pg 218]. This concludes the proof of Theorem 3. □

We note that these saddle-point distributions need not have finite support. However, in this case we have an equlibrium and with no further knowledge of each other's choice of strategy, the jammer and the quantizer should be content utilizing $dG^*(\theta)$ and $dF^*(z)$.

Using the Optimization Theorem and the Constrained Optimization Theorem we can derive necessary and sufficient conditions at these saddle points. Given any $dG(\theta)$ and the power constraint we define

$$\overline{U}_c(K_J, G) \triangleq \sup_{\substack{F \in \mathbf{S} \\ h_F \leq K_J}} - E_G(C(\theta, F)) \tag{28}$$

and given any $dF(z)$ we define

$$\overline{V}_c(F) \triangleq \sup_{G \in \mathcal{G}} E_G(C(\theta, F)) \tag{29}$$

where $\mathcal{G}$ is the space of distributions on Θ. Then we have

Theorem 6: The saddle-point strategies dF^*, dG^* satisfy to the following inequalities:

$$E_{G^*}(\int(-\bar{i}(z;\theta,F^*) - \lambda f(z))dF(z)) \leq E_{G^*}(-C(\theta,F^*)) - \lambda K_J \tag{30}$$

for some $\lambda \geq 0$, for all F where

$$\bar{i}(z;\theta,F) \triangleq \sum_{x,y} P(x)p(y|x,z,\theta)\log\frac{\int p(y|x,z,\theta)dF(z)}{\sum_x P(x)\int p(y|x,z,\theta)dF(z)}$$

Also

$$E_G(C(\theta,F^*)) \leq E_{G^*}(C(\theta,F^*)) \tag{31}$$

for all G.

Proof of Theorem 6:

For any F denote the weak derivative of $E_G(C(\theta,F))$ at G_0 as $D_{G_0}(E_G(C(\theta,F))$ and for any G denote the weak derivative of $E_G(C(\theta,F))$ at F_0 as $D_{F_0}(E_G(C(\theta,F))$. Using Lemma 3 and the Dominated Convergence Theorem, we have

$$D_{F_1}(E_G(-C(\theta,F_2)) = E_G(-\int\bar{i}(z;\theta,F_1)dF_2) + E_G(C(\theta,F_1)) \tag{32}$$

for any F_1, F_2.

Also

$$D_{G_1}(E_{G_2}(C(\theta,F) = E_{G_2}(C(\theta,F) - E_{G_1}(C(\theta,F). \tag{33}$$

Now letting $F_1 = F^*, G_1 = G^*$ in the first equation we have, using the Constrained Optimization Theorem and the Optimization Theorem and the properties of $E_G(C(\theta,F))$ as in Theorem 2, that a necessary and sufficient condition for F^* to achieve $\overline{U}_c(K_J,G^*)$ is

$$E_{G^*}(-\int(\bar{i}(z;\theta,F^*) - \lambda f(z))dF(z)) \leq E_{G^*}(-C(\theta,F^*)) - \lambda K_J \tag{34}$$

for some $\lambda \geq 0$, for all F.

Letting $F_1 = F^*, G_1 = G^*$ in the second equation gives us similarly that a necessary and sufficient condition to achieve $\overline{V}_c(F^*)$ is

$$E_G(C(\theta, F^*)) \leq E_{G^*}(C(\theta, F^*)) \tag{35}$$

for all G.

Since at a saddle-point $\overline{U}_c(K_J, G^*)$ and $\overline{V}_c(F^*)$ are simultaneously achieved, the theorem follows. □

4.1 The Remaining Cases

Case BII: Theorem 3 holds with $F(z)$ as a one-dimensional distribution.

Case CII: Although **S** is compact, it is not convex and so we cannot demonstrate that there is a saddle point strategy.

Case DII: Again we have that $E_G(C(\theta, F))$ is a Levy continuous functional of $dG(\theta)$ and is concave in $dG(\theta)$. Also $E_G(C(\theta, F))$ is Levy continuous in $(dF_1(z), \ldots, dF_D(z))$. However $E_G(C(\theta, F_1, \ldots F_D))$ is not convex in $(F_1, \ldots, F_D)$. Hence we cannot assert the existence of a saddle point in this case.

4.2 Fixed Quantizer

Before concluding this section we also point out that if we did not have randomized quantization then without "compatibility" the game would have a saddle-point where the jammer's saddle-point distribution need be concentrated at at most $M(L-1)+2$ points. We summarize this in Theorem 7.

Theorem 7: For any quantizer θ, there exists a pair of distributions $dP^*(x), dF^*(z)$ such that

$$I(\theta, P, F^*) \leq I(\theta, P^*, F^*) \leq I(\theta, P^*, F) \tag{36}$$

for all feasible dP, dF. Moreover $dF^*(z)$ can be chosen to be concentrated at at most $M(L-1)+2$ points and necessary and sufficient conditions for $dF^*(z)$ and $dP^*(x)$ are for some $\lambda_1, \lambda_2 \geq 0$

$$-i(z;\theta, F^*) \leq -I(\theta, P^*, F^*) + \lambda_1(f(z) - K_J) \tag{37}$$

for all $z \in K$ and

$$-i(z;\theta, F^*) = -I(\theta, P^*, F^*) + \lambda_1(f(z) - K_J) \tag{38}$$

for all $z \in E_0$ where $i(.;.,.)$ is as defined in Theorem 2 *with G concentrated* on θ. Also

$$I_x(\theta, P^*, F^*) = \lambda_2 \tag{39}$$

for all $x \ni P^*(x) > 0$ and

$$I_x(\theta, P^*, F^*) \leq \lambda_2 \tag{40}$$

for all $x \ni P^*(x) = 0$ where

$$I_x(\theta, P^*, F^*) \triangleq \sum_y p(y|x,\theta) \log \frac{p(y|x,\theta)}{\sum_x P^*(x) p(y|x,\theta)}.$$

Proof of Theorem 7:

From the proof of Theorem 5 we know that all we need to show is that $I(\theta, P, F)$ is (Levy) continuous in $dP(x)$. We show this by considering any sequence $dP_n(x) \xrightarrow{w}$

$dP(x)$ and showing $I(\theta, P_n, F) \to I(\theta, P, F)$. Since x belongs to the finite set A, weak convergence is equivalent to convergence in any finite-dimensional metric.

Now

$$|I(\theta, P_n, F) - I(\theta, P, F)| = |\sum_{x,y} P_n(x)p(y|x,\theta)\log\frac{p(y|x,\theta)}{\sum_x P_n(x)p(y|x,\theta)}$$

$$- \sum_{x,y} P(x)p(y|x,\theta)\log\frac{p(y|x,\theta)}{\sum_x P(x)p(y|x,\theta)}|$$

$$\leq \ |\sum_{x,y} P_n(x)p(y|x,\theta)\log\frac{p(y|x,\theta)}{\sum_x P_n(x)p(y|x,\theta)}$$

$$- \sum_{x,y} P_n(x)p(y|x,\theta)\log\frac{p(y|x,\theta)}{\sum_x P(x)p(y|x,\theta)}|$$

$$+ |\sum_{x,y} P_n(x)p(y|x,\theta)\log\frac{p(y|x,\theta)}{\sum_x P(x)p(y|x,\theta)}$$

$$- \sum_{x,y} P(x)p(y|x,\theta)\log\frac{p(y|x,\theta)}{\sum_x P(x)p(y|x,\theta)}|$$

$$\leq |\sum_{x,y} P_n(x)p(y|x,\theta)||\log\frac{\sum_x P_n(x)p(y|x,\theta)}{\sum_x P(x)p(y|x,\theta)}|$$

$$+ \sum_x D|P_n(x) - P(x)| \tag{41}$$

where $D = \max_{x,y} p(y|x,\theta)\log\frac{p(y|x,\theta)}{\sum_x p(y|x,\theta)}$

$$\leq LD|\log\frac{\sum_x P_n(x)p(y|x,\theta)}{\sum_x P(x)p(y|x,\theta)}|$$

$$+ \sum_x D|P_n(x) - P(x)|. \tag{42}$$

Again since A is finite we can say that for all $\delta > 0$ $\exists N$ such that for all n > N

$$1 - \delta \ \leq \frac{P_n(x)}{P(x)} \leq 1 + \delta \ \ \forall x \in A$$

$$1 - \delta \ \leq \frac{P_n(x)p(y|x,\theta)}{P(x)p(y|x,\theta)} \leq 1 + \delta \ \ \forall x \in A$$

$$1-\delta \ \leq \frac{\sum_x P_n(x)p(y|x,\theta)}{\sum_x P(x)p(y|x,\theta)} \leq 1+\delta \ \ \forall x \in A. \tag{43}$$

By the continuity of the log function we can say that $\forall \epsilon > 0 \ \ \exists \delta > 0 \ni$

$$-\epsilon \leq |\log \frac{\sum_x P_n(x)p(y|x,\theta)}{\sum_x P(x)p(y|x,\theta)}| \leq \epsilon.$$

The second term in (41) can also clearly be made $\leq \epsilon$ for sufficiently large n. Thus the continuity of $I(\theta, P, F)$ w.r.t. P is confirmed and the first part of the theorem follows. The bound on the number of points of support of dF^* follows from Theorem 1(a). The necessary and sufficient conditions are derived as before from Theorem 3 and well-known results about channel capacity [Gall 68, pg.91]. □

5 Conclusions

We have constructed fairly general channel models which are capable of representing a number of jamming situations. The jammers we have considered have all been non-adaptive and using results from the compound channel we are able to give operational significance to our minimax performance measures,i.e. we can assert the existence of encoders and decoders which can perform at arbitrarily low probabilities of error at rates close to our performance measures. Our analysis is also clearly applicable to many restrictions on the jammer's strategy set other than the ones we have considered.

In the case with the decoder uninformed (case I) we have shown that the worst-case jammer strategy (as well as best communicator strategy) needs only be one of the class of distributions with finite support. We have a bound on the number of these points of support in terms of the sizes of the input and the output alphabet. Thus we have reduced the computation of the worst case jamming strategies to a finite-dimensional non-linear programming problem. Moreover we can characterize these distributions by necessary and sufficient conditions which are fairly easy to test.

In the cases with the decoder informed we reduce the communicator's strategy set (either by using the "compatibility" assumption or by fixing a quantizer) . In this case when we have convexity with respect to the jammer's strategy (as in cases AII and BII) we are able to demonstrate the existence of a saddle-point strategy. For the case with non-randomized quantization we are further able to characterize these saddle-point strategies using the earlier theory.

As we have mentioned earlier all the above presupposes non-adaptive jamming. The compound channel model which we use indirectly by our choice of objective function is appropriate to use in this case. We can allow for more sophisticated jammers if we incorporate the cases where the jammer's strategies are allowed to depend on the previous (and present) channel inputs. The appropriate channel model to use then is that of the arbitrarily "star" varying channel (A^*VC) [Csiz 81, pg.233]. This model generalizes the arbitrarily varying channel (AVC) and includes it as a special case. It is known that the m-capacity (i.e. capacity with maximum probability of error over all the codewords) of the A^*VC is the same as that of the corresponding AVC [Csiz 81, pg.232]. This capacity is known for the case of binary output alphabet (and finite input alphabet) and is known to be equal to $\max_{dP(x)} \min_{W \in \overline{\overline{\mathcal{W}}}} I(X;Y)$ where X and Y are the input and the output respectively, W is any channel chosen from the set of channels $\mathcal{W}$ and $\overline{\overline{\mathcal{W}}}$ is the row-convex closure of $\mathcal{W}$ [Csiz 81]. In our case the jammer's strategy set corresponding is already row-convex closed and hence the appropriate programs would be

a) For the communicator:

$$\max_{(dG(\theta), dP(x))} \min_{dF(z)} I(G, F)$$

b) For the jammer

$$\min_{dF(z)} \max_{(dG(\theta), dP(x))} I(G, F)$$

which is the same objective function as we have used. Similarly, in the case with decoder informed we would obtain the same objective functions. Thus, all the results derived in the previous chapter for the case of mutual information can be extended to the case of the A^*VC channel with binary output. This model may be viewed as a worst-case representation of adaptive jamming. Unfortunately the m-capacity of the AVC is as yet unknown for output sizes greater than 2. On the other hand the a-capacity of the AVC (i.e. the capacity with average probability of error) is known to be either 0 or else $\max_{dP(x)} \min_{W \in \overline{\mathcal{W}}} I(X;Y)$ where $\overline{\mathcal{W}}$ is the convex closure of the set $\mathcal{W}$ to which W belongs [Csiz 81, pg.214]. Since in our model the set of channels is convex as well as row-convex the a-capacity is known to be greater than 0 iff the m-capacity is greater than 0 [Ahls 78]. Thus with average probability of error whenever the jammer's strategy set is such that he cannot force the capacity to be 0 then all the results of the preceding chapter extend to the case of the A^*VC channel.

Appendix A

Lemma 3 : $I'_{F_1}(G;F_2) = \int i(z;G,F_1)dF_2(z) - I(G;F_1)$

where $i(z;G,F_1) = \sum_{x,y} p(x)p(y|x,z)\log\left(\frac{\int p(y|x,z)dF_1}{\sum_x p(x)\int p(y|x,z)dF_1}\right)$.

Proof of Lemma 3 :

$$I'_{F_1}(G;F_2) = \lim_{\alpha\downarrow 0}\frac{1}{\alpha}\left\{\sum_{x,y} p(x)\left(\int\int p(y|x,z,\theta)[(1-\alpha)dF_1+\alpha dF_2]dG(\theta)\right).\right.$$

$$\log\frac{(\int\int p(y|x,z,\theta)[(1-\alpha)dF_1+\alpha dF_2]dG(\theta))}{\sum_x p(x)(\int\int p(y|x,z,\theta)[(1-\alpha)dF_1+\alpha dF_2]dG(\theta)}$$

$$-\sum_{x,y} p(x)(\int\int p(y|x,z,\theta)dF_1 dG(\theta)).$$

$$\left.\log\frac{(\int\int\ p(y|x,z,\theta)dF_1dG(\theta))}{\sum_x p(x)(\int\int p(y|x,z,\theta)dF_1dG(\theta)}\right\}. \tag{44}$$

Denoting $\int p(y|x,z,\theta)dG(\theta)$ as $p(y|x,z)$

$$I'_{F_1}(G;F_2) = \lim_{\alpha\downarrow 0}\frac{1}{\alpha}\left\{\sum_{x,y} p(x)\left[\int p(y|x,z)[(1-\alpha)dF_1 + \alpha dF_2]\right.\right.$$

$$\log\frac{\int p(y|x,z)[(1-\alpha)dF_1+\alpha\ dF_2]}{\sum_x p(x)\int p(y|x,z)[(1-\alpha)dF_1+\ \alpha dF_2]}$$

$$\left.-\int p(y|x,z)dF_1\log\frac{\int p(y|x,z)dF_1}{\sum_x p(x)\int\ p(y|x,z)dF_1}\right\}$$

$$= \lim_{\alpha\downarrow 0}\frac{1}{\alpha}\left\{\sum_{x,y} p(x)\left[\int p(y|x,z)\alpha dF_2\log\left(\frac{p(y|x,z)[(1-\alpha)dF_1+\alpha dF_2]}{\sum_x p(x)\int p(y|x,z)[(1-\alpha)dF_1+\alpha\ dF_2]}\right)\right.\right.$$

$$\left.\left.-\int p(y|x,z)\alpha dF_1\log\left(\frac{\int p(y|x,z)[(1-\alpha)dF_1+\alpha dF_2]}{\sum_x p(x)\int p(y|x,z)[(1-\alpha)dF_1+\alpha dF_2]}\right)\right]\right\}$$

$$+\lim_{\alpha\downarrow 0}\frac{1}{\alpha}\sum_{x,y}p(x)\left[\int p(y|x,z)dF_1\log\left(\frac{\int p(y|x,z)[(1-\alpha)dF_1+\alpha dF_2]}{\sum_x p(x)\int p(y|x,z)[(1-\alpha)dF_1+\alpha dF_2]}\right)\right.$$

$$\left.-\int p(y|x,z)dF_1\log\left(\frac{\int p(y|x,z)dF_1}{\sum_x p(x)\int p(y|x,z)dF_1}\right)\right]$$

$= a + b(\text{say})$.

By choosing a sequence $\alpha_n \downarrow 0$ and using weak convergence of $(1-\alpha_n)dF_1 + \alpha_n dF_2$ to dF_1

$$a \;=\; \int i(z;G,F_1)dF_2 - I(G;F_1)$$

$$b=\frac{d}{d\alpha}\left[\sum_{x,y}p(x)\int\; p(y|x,z)dF_1\log\left(\frac{\int p(y|x,z)[(1-\alpha)dF_1+\alpha dF_2]}{\sum_x p(x)\int p(y|x,z)[(1-\alpha)dF_1+\alpha dF_2}\right)\right]_{\alpha=0}$$

at $\alpha = 0$.

Taking the derivative

$$b=\sum_{x',y}p(x')\int p(y|x',z)dF_1\left\{\frac{\sum_x p(x)\int p(y|x',z)[(1-\alpha)dF_1+\alpha dF_2]}{\int p(y|x,z)[(1-\alpha)dF_1+\alpha dF_2]}\right.$$

$$\cdot\frac{1}{d^2}\left[\left(\sum_x p(x)\int p(y|x,z)[(1-\alpha)dF_1+\alpha dF_2]\right)\int p(y|x',z)(dF_2-dF_1)\right.$$

$$\left.\left.-\int p(y|x',z)[(1-\alpha)dF_1+\alpha dF_2]\left(\sum_x\; p(x)\int p(y|x,z)(dF_2-dF_1)\right)\right]\right\}$$

where $d \triangleq \sum_x p(x)\int p(y|x,z)[(1-\alpha)dF_1+\alpha dF_2$

After some algebraic manipulation it can be shown that $b \to 0$ as $\alpha \downarrow 0$.

Appendix B

Here we consider a communication game with two players, player A who chooses an input distribution r on the M-ary input alphabet, and player B who chooses the $M \times L$ transition probability matrix. Let X and Y denote the input and output random variables respectively and let n_i denote the distribution of the random variable associated with the conditional density $p(y|x_i)$. Let the set of all feasible $\underline{n}$'s $(= (n_1, \ldots, n_M))$ be compact. The channel $p(y|x)$ is a function of $\underline{n}$ $(= (n_1, \ldots, n_M))$. Assume this function is linear and that for a choice of $n_i = n, \ i = 1, \ldots, M$ the channel chosen is symmetric. Let $I(r, \underline{n}) \triangleq I(X;Y)$ when A's choice is r and B's choice is $\underline{n}$. Let $n_1, \ldots, n_M$ be constrained by $f_i(n_1, \ldots, n_M) \le c_i$ $i = 1, \ldots, c$ where f_i is a convex, symmetric function of $n_1, \ldots, n_M$, i.e. f_i is invariant under any permutation of $n_1, \ldots, n_M$. Then a saddle-point strategy exists for both players and for player A it is to chooose a uniform distribution on the input and for player B it is to choose all the components of $\underline{n}$ equal, i.e. there exists $\underline{n}^*$ with all its components equal such that

$$I(r, \underline{n}^*) \le I(r^*, \underline{n}^*) \le I(r^*, \underline{n})$$

where r^* corresponds to the uniform input distribution.

Proof: Step 1: $I(r, \underline{n}^*) \le I(r^*, \underline{n}^*)$

This follows from the fact that the mutual information between the input and the output of a symmetric channnel is maximized by the uniform distribution.

Step 2: $I(r^*, \underline{n}^*) \le I(r^*, \underline{n})$

Since $I(X;Y)$ is a convex function of $p(y|x)$ which is linear in $\underline{n}$, $I(r, \underline{n})$ is convex in $\underline{n}$. Moreover, given the form of the constraints the set of feasible $\underline{n}$'s is a convex set.

Now for any $\epsilon > 0$ let $\inf I(r^*, \underline{n}) + \epsilon$ be achieved at some $\underline{n}_1 \neq \underline{n}^*$. Then we show $I(r^*, \underline{n}^*) \leq I(r^*, \underline{n}_1)$ proving that the minimum is also achieved at $\underline{n}^*$. The use of a uniform distribution on the input and the symmetry of the constraints implies that for any permutation of $\underline{n}_1$ ($\underline{n}_1^\alpha$ say) we have a new channel $p^\alpha(y|x)$ which involves just a relabelling of the inputs of the original channel. The mutual information $I(r^*, \underline{n}_1)$ is equal to $I(r^*, \underline{n}_1^\alpha)$. Now consider all the $M!$ permutations of $\underline{n}_1 = \underline{n}_1^\alpha$: $\alpha \in T$ (all the permutations are not distinct but it does not matter). Take the convex combination $\frac{1}{M!} \sum_{\alpha \in T} \underline{n}_1^\alpha = \underline{n}_e$(say). Every component of $\underline{n}_e$ is equal to $\frac{1}{M!} \sum_{i1}^{M} n_{1_i}$. Also from the convexity of $I(r^*, \underline{n})$ w.r.t. $\underline{n}$ we know that

$$\begin{aligned} I(r^*, \frac{1}{M!} \sum_{\alpha \in T} \underline{n}_1^\alpha) &\leq \frac{1}{M!} \sum_{\alpha \in T} I(r^*, \underline{n}_1^\alpha) \\ &= I(r^*, n_1) \end{aligned}$$

Therefore

$$I(r^*, \underline{n}_e) \leq I(r^*, \underline{n}_1)$$

and hence $\inf I(r^*, \underline{n}) + \epsilon$ is achieved at $\underline{n}_e$ too. The result then follows from the observation that $I(r^*, \underline{n})$ is concave in r.

References

[1] [Ahls 78] R. Ahlswede, "Elimination of correlation in random codes for arbitrarily varying channels", *Zeitschrift fur Wahrscheinlichkeitstheorie*, no. 33, pp. 159-175, 1978

[2] [Ash 65] R.Ash, *Information Theory*, Interscience Publishers, 1965

[3] [Ash 72] R. Ash, *Real Analysis and Probability*, Academic Press,Inc., 1972

[4] [Aubi 82] J.P. Aubin, *Mathematical Methods of Game and Economic Theory*, North-Holland, 1982

[5] [Blac 57] N.M. Blachman,"Communication as a game", *Wescon 1957 Conference Record*, 1957

[6] [Blac 54] D. Blackwell, M.A. Girshick, *Theory of Games and Statistical Decisions*, Dover Publications, INc., 1954

[7] [Blac 59] D. Blackwell, L. Breiman, A.J. Thomasian,"The capacity of a class of channels",*Annals of Mathematical Statistics*, no.30, pp.1229-1241

[8] [Blac 60] D. Blackwell, L. Breiman, A.J. Thomasian,"The capacities of certain channel classes under random coding", *Annals of Mathematical Statistics*, no.31, pp.558-567

[9] [Bord 85] J.M. Borden, D.J.Mason, R.J.McEliece,"Some information theoretic saddlepoints", *SIAM Journal on Control and Optimization*, vol. 23, no. 1, Jan 1985

[10] [Chan 85] L.F. Chang, An Information-Theoretic Study of Ratio-Threshold Antijam Techniques, Ph.D. Thesis, University of Illinois at Urbana-Champaign, 1985

[11] [Csiz 81] I. Csiszar and J. Korner, *Information Theory : Coding Theory for Discrete Memoryless Systems*, Academic Press, 1981

[12] [Dobr 59] R.L. Dobrushin,"Optimum information transmission through a channel with unknown parameters", *Radio Engineering Electronics*, vol. 4, no. 12, 1959

[13] [Dubi 62] L.E. Dubins,"On extreme points of convex sets", Journal of Mathematical Analysis and Applications, pp. 237-244, 1962

[14] [Eric 85] T. Ericson,"The arbitrarily varying channel and the jamming problem", *Internal Report LiTH-ISY-I-0772, Department of Electrical Engineering, Linkoping University, Sweden*, 1985

[15] [Gall 68] R.G. Gallagher, *Information Theory and Reliable Communication*, Wiley, 1968

[16] [Hegd 87] M.V. Hegde, Performance Analysis of Coded, Frequency-Hopped Spread-Spectrum Systems, Ph.D. Thesis, University of Michi-

gan, Aug. 1987

[17] [Karl 59] S. Karlin, *Mathematical Methods and Theory in Games, Programming and Economics*, vols 1 and 2, Addison- Wesley, 1959

[18] [Loev 77] M. Loeve, *Probability Theory I*, Springer-Verlag, 1977

[19] [Luen 69] D.G. Luenberger *Optimization by Vector Space Methods*, Wiley, 1969

[20] [Ma 84] H.H. Ma, M.A. Poole,"Error-correcting codes against the worst-case partial-band jammer", *IEEE Transactions on Communications*, vol. 32, pp.124-133, Feb. 1984

[21] [McEl 77] R.J. Mceliece, *The Theory of Information and Coding*, Addison-Wesley, 1977

[22] [McEl 84] R.J. McEliece, W.E. Stark,"Channels with block interference", *IEEE Transactions on Information Theory*, vol. 30, pp.44-53, Jan. 1984

[23] [McEl 83] R.J. McEliece, E.R. Rodemich,"A study of optimal abstract jamming strategies vs. noncoherent MFSK", *Military Communications Conference Record, 1983*, pp. 1.1.1 -1.1.6, 1983

[24] [McEl 83] R.J. McEliece,"Communication in the presence of jamming-an information theoretic approach", in *Secure Digital Communications*, Springer- Verlag, pp. 127-166 1983

[25] [McEl 82] R.J. McEliece and W.E Stark,"The optimal code rate vs. a partial band jammer", *Milcom Record 1982* , pp. 45.3.1 - 45.3.5, 1982

[26] [Peng 86] W.C. Peng, Some Communication Jamming Games, Ph.D. Thesis, University of Southern California, Jan 1986

[27] [Root 61] W.L. Root,"Communication through unspecified additive noise", *Information and Control*, vol. 4, pp. 15-29, 1961

[28] [Scha 68] H. Schaubert, *Topology*, Macdonald and co. Ltd., 1968

[29] [Simo 85] M.K. Simon, J.K. Omura, R.A. Scholz, B.K. Levitt *Spread Spectrum Communications, vols 1,2 and 3*, Computer Science Press, 1985

[30] [Smit 71] J.G. Smith,"The information capacity of amplitude- and variance-constrained scalar gaussian channels", *Information and Control*, vol. 18, pp.203-219, 1971

[31] [Star 82] W.E. Stark, Coding for Frequency-Hopped Spread-Spectrum Channels with Partial-band Interference, Ph.D. Thesis, University of Illinois at Urbana-Champaign, 1982

[32] [Star 85a] W.E. Stark,"Coding for frequency-hopped spread-spectrum communication with partial-band interference-Part 1: capacity and cut-off rate", *IEEE Transactions on Communications* vol. 33, no. 10, Oct. 1986

[33] [Star 85b] W.E. Stark,"Coding for frequency-hopped spread-spectrum communication with partial-band interference-Part 2: Coded performance", *IEEE Transactions on Communications* vol. 33, no. 10, Oct. 1986

[34] [Star 86] W.E. Stark,D. Teneketzis, S.K. Park,"Worst-case analysis of partial-band interference", *Proceedings of the 1986 Conference on Information Sciences and Systems* , 1986

[35] [Stig 66] I.G. Stiglitz,"Coding for a class of unknown channels", *IEEE Transactions on Information Theory*, vol. 12, pp.189-195, 1966

[36] [Vite 79] A.J. Viterbi, J.K. Omura, *Principles of Digital Communication and Coding*, McGraw-hill, 1979

[37] [Wits 80] H.S. Witsenhausen," Some aspects of convexity useful in Informatios theory', *IEEE Transactions on Information Theory*, vol. 26, pp.265-271, May 1980

[38] [Wolf 78] J.Wolfowitz, *Coding Theorems of Information Theory*, Springer-Verlag, 1978

END
DATE
FILMED
5-88
DTIC